图书情报工作
第63卷 第1期 2019年1月

图书馆智慧服务中的数据获取*

李玉海　朱泽　夏红玉
华中师范大学信息管理学院　武汉 430079

摘要：［目的/意义］定义图书馆智慧服务的内涵，建构支持智慧服务的数据属性模型，并进行方法和过程的讨论，从而促进智慧服务的实现。［方法/过程］对图书馆智慧服务的含义进行解释，提出图书馆服务的智慧程度和数据的交互性、交互频度、响应时间以及数据感知深度的关系模型，论述不同数据属性对智慧服务的影响，并提出数据获取的方式。［结果/结论］数据是现代图书馆实现智慧服务的必要支撑，图书馆必须依靠足够的技巧与耐心，挖掘用户的潜在数据，并基于相应的数据交互实现快速、深入、灵活的智慧服务。

关键词：智慧图书馆　智慧服务　数据获取　数据交互
分类号：G250
DOI：10.13266/j.issn.0252-3116.2019.01.005

1 引言

近年来，随着信息技术的飞速发展及应用，以“互联网+”为代表的创新思维对图书馆这个传统的文献服务体系提出了新的挑战。当前图书馆的服务内容由曾经的图书、期刊文献借还、阅览等内容，扩展到了文献传递、数据比对分析、个性化定制等诸多项目。图书馆的服务内容不再是单纯的对文献的存储和管理，而是更加灵活地对数据进行挖掘，以贴近用户的需求。这种变化具体表现在馆藏资源的配置更加科学，资源的揭示更加深刻、全面，资源的发现更加准确、快捷。

一方面，无论是图书馆服务的提供者还是接受者，更多地感受到图书馆仿佛具有人的智慧，知道人们想做什么、需要什么，在减少人工服务和冗余交流的前提下，尽量感知并满足人们的要求。人类的需求表达过程越来越轻松，也越来越准确，这些现象背后都是数据的挖掘、应用和运行在发挥着作用。另一方面，网络技术和自动化设备在图书馆的应用，客观上也产生了大量数据。例如装有时间、位置、触摸操控的传感器以及能记录用户上网行为的移动终端、服务系统软件，可以贡献图书馆环境、文献资源的客观状态数据，以及用户的实时、原始行为数据，还有人们利用软件工具对客观数据的加工处理所产生的二次数据，相比模糊状态下的一次数据，这种数据更具生命力和应用价值。正是对这类数据的挖掘和运用，才使得图书馆应用系统具有智慧服务功能。因此，对数据获取的研究将是提升服务智慧程度的重要途径。笔者将对近年着眼于图书馆智慧服务的相关研究进行概述，梳理相关的研究脉络，并从中找到相应的突破口，对图书馆未来智慧服务所需要的关键技术和可能的走向展开探讨。

2 文献综述

对智慧图书馆的研究逐渐成为图书馆学界的热门课题，相关的发表文献数量逐年增加。周承聪等[1]通过分析1999-2015年时间段内发表的相关论文，得出了我国智慧图书馆的相关研究还处于起步阶段，论文和关注量较少，通过对文献的关键词和高频词进行分析后，认为我国智慧图书馆的相关研究主要集中在6个方面：智慧图书馆的技术环境；信息技术的实际应用；社会环境；智慧服务；个性化服务；以及数字图书馆建设。而其中的智慧服务由于依赖主流信息技术的发展，研究趋势不够清晰。

早期对图书馆智慧服务影响较大的是技术理念。严栋[2]于2010年就提到了物联网对于智慧图书馆服务能力的塑造作用，认为传感、传输和应用控制网络的应用将全面提升图书馆的感应、传递等能力。虽然如

* 本文系中央高校基本科研业务费重大培育项目“智慧图书馆管理系统关键技术与应用研究”（项目编号：CCNU18JCXK04）研究成果之一。
作者简介：李玉海（ORCID：0000-0003-2256-0183），院长，教授，博士生导师；朱泽（ORCID：0000-0003-1897-886X），硕士研究生，通讯作者，E-mail：1074908864@qq.com；夏红玉（ORCID：0000-0002-8995-8519），助理馆员，硕士研究生。
收稿日期：2018-09-30　修回日期：2018-11-18　本文起止页码：30-35　本文责任编辑：徐健

今图书馆的发展确实与物联网联系越发密切,但整体来看研究的进展比较缓慢;文献中体现出比较主流的研究方向是将信息技术理论实际应用到图书馆的智慧服务中,强调实用性;M. Buscema 等[3]将人工神经网络引入智慧图书馆建设,通过对学生的能力和偏好的整合设计图书馆的服务流程。陈臣[4]提出基于读者阅读习惯和体验的大数据,融合智慧服务为读者打造个性化智慧阅读服务;曾子明等[5]融合了情景感知对图书馆读者的情景信息进行挖掘和处理,从而建立相关的个性化服务模型,进行智慧服务个性化的探索。

另外,部分学者通过对图书馆环境和读者需求进行分析,并针对性地提供相应的智慧服务。陈远等[6]结合泛在信息环境提出了泛在智慧服务智慧图书馆层次模型,构建了一个以信息技术为支撑,面向用户泛在需求的未来智慧图书馆形态;刘宝瑞等[7]借助新智能技术群重构了用户体验,设想智慧服务对用户体验的影响,以及未来智慧图书馆服务体验的变化。

总体来看,关于智慧服务的研究多为框架的构想和设计,但对智慧服务具体的实现路径和技术运用的研究却很少,一方面是对智慧图书馆整体研究分散,另一方面也是缺乏核心信息技术的支持,对于用户需求的分析过于缥缈,找不到着力点。

笔者试图从数据获取的角度,创造性地对图书馆应该如何挖掘用户需求数据进行讨论,并借助多元化的技术设备,为实现智能化图书馆的智慧服务和数据获取提出可能性。

3 图书馆智慧服务的内涵

在讨论图书馆如何实现智慧服务之前,需要先明确图书馆智慧服务的内涵。从狭义上来看,图书馆智慧服务即是图书馆应该能够具备类人的智慧,用以感知用户的需求,并及时有效地解决问题。这种类人的服务体验需要图书馆时刻伴随在用户身边。当用户需要图书馆服务时,无论是一本书的借阅或者是某类研究项目的国内外成果、进展情况,还是文献资源配置策略等,只要通过移动终端,以语音、文本、手势等发送书名、作者、项目名称及关键词等不同类型的信息需求时,图书馆会立即对用户的请求加以响应,并且突破时间和空间的约束,将最终符合用户需要的结果传递给他们。对于实现这类智慧服务的过程,我们可以仿照人类解决问题的基本做法,总结为感知交互、思考判断和决策执行三个环节。因此,从整体上看,图书馆的智慧服务应该包含以下方面:

3.1 泛在服务环境

泛在服务环境是实现智慧服务的基本条件,保障数据的交换传输,是智慧服务的组成部分之一。泛在服务环境主要指无所不在的网络接入条件,将用户所在地纳入图书馆的服务范围,使用户能随时随地在需要服务时,向图书馆发出请求。在当今移动网络时代,高速网络的铺设已经成为一项城市建设的基础工程项目。而伴随着5G通信技术的普及应用,万物互联将得以实现,网络的覆盖面将会更大,数据传输速度更快。打破空间和地域局限的图书馆服务,使得用户对图书馆的泛在服务体验会更满意。

3.2 感知交互服务

对于服务接受者作为图书馆服务的客体,是智慧服务的目标群体,其需求为图书馆服务提供动力和目标。但同时由于用户本身受各类客观或主观因素的影响(如用户的知识结构、工作环境、生活环境、文化背景等),其需求有层次之分(见图1),也就是说多数时候用户对服务需求的表达处在一个混沌和模糊的状态。这时,图书馆的主动感知交互能力就显得尤其重要。

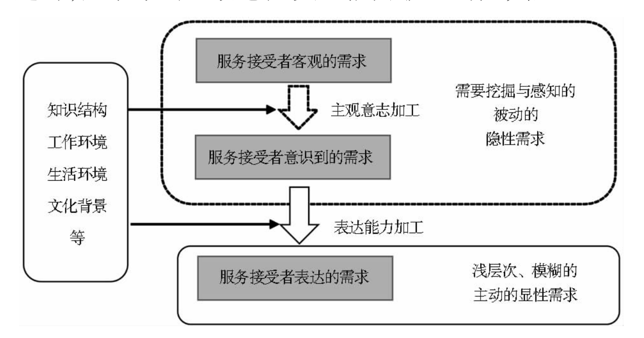


图1 服务接受者的需求层次

感知交互服务是实现智慧服务的第一步,感知往往是隐形的,交互一般是显形的。图书馆管理服务系统会根据用户开放的移动终端设备的感应数据和历史上网行为数据,自动感知用户的状态,预测下一步的行动。而交互则是满足用户向系统发出实时请求的需要,且问题较多或请求表述不明确,需要不断提问、反复确认的情况下,系统启动问答交互功能,在人工沟通或者智能回复的情况下,尽可能准确理解请求意图,在这一过程中同时获取数据并进行数据更新,以便为后续处理环节提供输入信息,最后实现最精确的智慧服务。

这一过程对服务接受者而言,往往是在其不知晓的情况下,以不侵犯个人隐私为前提下发生的,用户只需按照系统提供的服务功能享受服务,如查看推送的资料、打开感兴趣的内容、更新自己的资料等。而对服

务提供者来说，通过更敏锐的设备和更灵活的服务决策流程，为用户提供更及时、有效和个性化的精准服务。人工服务将更多地作为流程的监控者和终端反馈的处理者，这样的转变对于服务提供者来说，已经从繁琐的咨询业务中剥离出来，富有知识服务的工作不仅意味着更专业、准确的职能定位，更意味着高效、精确的服务输出。

3.3 机器学习

机器学习是伴随着人工智能的发展所提出的理念，从原理上来看，机器学习就是依靠能力和数据储备的提升，实现、强化机器的自我知识获取能力[9]。当前机器学习的两种主要方法包括：自身学习和模仿学习，通过对机器逻辑语言的设计或者对人类学习行为的模仿，使机器实现类人的演绎、归纳和类比等延伸数据内涵的行为。在图书馆领域，大量的文献数据和用户数据交叉，仅凭人力梳理就已经捉襟见肘，更难以实现数据深层的挖掘，而拥有高速的计算能力、巨大的存储空间和灵活适应的机器学习，将为图书馆的海量数据激活与应用提供可能。例如高效的机器学习在收集到大量的在校学生阅读和图书馆活动记录后，即可通过归纳和演绎整合出具体的"阅读圈"，从而便于图书馆进行有引导性、针对性的服务和阅读推广活动等。可以说在巨大的数据体量下，机器学习弥补了人工的不足，使得数据分析处理以及后续服务更加智能化。

3.4 决策执行

在现有图书馆服务系统中，做出决策和执行的往往是服务者；但人工所存在的响应滞缓、分工限制等问题，常常会导致决策和执行力较为低下、效率不高。而智慧服务系统不存在服务间歇和知识盲区，根据程序逻辑设定和机器判断，对用户的需求进行思考判断，在跳过人为服务的情况下，自主进行决策执行，满足用户的需要。例如在商业领域已经得到广泛运用的自动回复功能，根据用户提出的问题进行关键词识别，在知识库中获取相关的答案或解决方案，并推送给提问者，还可以通过不断地交流，精确推送结果，实现用户的自助咨询服务。在一定程度上，服务系统的自行决策执行，将极大地缓解人工重复服务的繁重劳动，同时提升服务效率和速率。

4 智慧服务的数据属性

从上述定义中能够看到，智慧服务定位的核心是用户需求，而用户需求的来源则是对数据的分析。可以说，数据是打通智慧图书馆与用户连接的关键"钥匙"，是图书馆智慧服务的源泉，图书馆用户从提出需求到上网操作过程中产生的行为数据，都可以作为提升图书馆服务质量的依据，而如何准确、及时地获取数据则成为了人们关注的焦点。

4.1 数据属性模型

事实上，早期的被动接受数据的服务模式已经远不能适应智能时代人们的要求。在追求人力解放、服务分工细化的今天，主动掌握用户数据，实现数据交互至关重要。从单向的需求提交到双向的数据交互，现在已经发展到对更深层的数据进行挖掘利用。一般而言，对现有各类数据特性分析，可以将支持智慧服务的数据属性模型表示为：

$$S = f(d_i, D_f, D_r, D_d) \qquad \text{公式(1)}$$

在公式(1)中，D_i 表示数据交互性，D_f 是数据更新频度，D_r 是响应速度，D_d 数据揭示深度(细粒度)。公式(1)说明，图书馆的智慧服务程度 S 与数据的交互性、更新频率、响应速度以及数据揭示深度有关，至于是一种什么样的函数关系 f，则需要通过大量实验探讨。

4.2 数据属性与智慧服务的关系

公式(1)中的各项数据属性都与智慧服务的质量密切相关。其中，数据交互性 D_i 决定了数据的获得渠道、种类和数量，直接决定了智慧服务的数据质量。交互性越强，相当于感知越灵敏，产生的有意义的数据也就越多，交互双方相互之间的了解越充分，带来的效果是提供的服务越有价值。

数据更新频度 D_f 是单位时间内数据更新的次数。它决定了智慧服务的准确度和时效性。当图书馆服务端的数据更新时间太慢，落后于用户提交请求和等待结果的时间，智慧服务就会显得过时和陈旧，同时，数据更新频度会带来数据采集少，使计算判断结果不准确，最终导致服务失去价值而显得冗余。

数据的响应速度 D_r 是指在数据交互过程中，某一方发出请求后，对方给出回应的时间。时间越短，响应越快，说明智慧服务的活力强、效率高，能大大提升工作效率，提高服务满意度。迟钝的数据响应将导致大量的数据流失，丧失数据源的信任，从而降低使用者对智慧服务的期待。

数据的揭示深度 D_d 是指对一篇文献中对语义揭示的程度，即细粒度。例如目前对一篇论文的揭示只到关键词、作者等，还不能从语义角度识别一句话。图书也是如此，只能是 MARC 数据，很难深入到任意一页的任意一句。揭示深度决定了智慧服务的个性化程

李玉海，朱泽，夏红玉. 图书馆智慧服务中的数据获取[J]. 图书情报工作,2019,63(1):30－35.

度,系统所挖掘的数据细粒度越高,就越能将整篇文献内容碎片化,针对数据的处理和延展服务就越有针对性。

因此,当图书馆服务者与用户之间的数据交流所得到的数据属性能够实现该模型的正向有利增长时,图书馆便能提供更精确、实时、高效、多样化、个性化的服务,也就是我们所追求的智慧服务。

5 数据的获取

从智慧服务的数据属性模型可以看出:影响智慧服务的程度和质量的首先是数据是否具有交互性,然后是交互的程度。下面从数据是否交互以及交互程度来探讨数据的获取方法。

5.1 非交互性数据获取

根据数据属性模型,非交互性数据就是不需要人与服务系统的交流,由系统设备、非交互传感器及软件运行而单向感知、提交的数据,如系统日志、图书馆的温度、湿度、地理位置、发布的广播信息等。非交互性数据大多是事实数据,客观性强,是智慧服务中机器学习、思考决策的重要参考依据。

非交互性数据的获取方式主要通过设备和软件系统,对目标对象进行识别和感知,进而通过自身的应用设置实现数据的感知和获取。在社会生活领域,这类功能的传感器包括:室内湿度传感器、光敏传感器、运动类的加速度(重力)传感器等,在图书馆使用较多的RFID。随着这些传感器的广泛运用,大众用户能感受到无形的"服务者"更贴心了,这也体现了这类传感器获取数据的单向性特点,用户本身无法察觉到自身数据被采集了,而传感器也是被动感应目标数据,与用户间不存在交互性。

提高此类数据获取的广度和精度一般采取以下方法:①多布设数据采集点,增加传感器数量,如增加不同方向的场景摄像头,缩短光线、声音感应器安装的距离等;②提高传感器和设备的感应精度和灵敏度,实现数据的高精度控制;③优化选择信号采样频率,提高数据获取的实时性。

5.2 交互性数据获取

对于重视用户体验与反馈的智慧服务来说,缺乏交互性的数据不足以形成用户满意的服务体系。而在互联网络空前发达的当下,交互性的数据交流、获取正在通过各种形式得以实现,如较为普遍的 BBS、微博、微信、微信公众平台等都能成为交互性数据的获取渠道,尤其是对于图书馆用户这类受众较易于形成网络社群,或信息接受能力较强的群体而言,活跃度较高的网络渠道将成为其与图书馆之间的主要数据交互场所。

这类数据的获取带有明显的交互性、即时性,交流双方既可延时也可即时在网络上进行沟通、交流。这一类的交互数据是参与交流的双方的情感、情绪、思维过程的文本表达,具有挖掘价值。由于网络的即时性特点,当交互的频度 D_f 越高时,服务的智慧程度越高,交互过程中对请求方响应 D_r 越快,即响应时间越短,服务满意度就越好。对于这类的数据获取,有效的方式是提升交互环境质量,如多提供可供选择的问题库、回复消息列表、表情图标,设置服务者的引导功能增强交互,以及通过系统设置和提示通知等方式加快交互响应过程,增强响应处理能力。

5.3 深度交互数据的获取

当前信息技术可以使我们设置应用设备,如手机、kindle 等的功能来捕获、跟踪用户在整个文献阅读过程中的关注热点,从而感知其感兴趣的内容,也就是数据获取的深度。这是实现智慧服务的重要方向,以成为学界、业界乃至企业普遍关注的问题。

深度交互的数据获取可以通过人机交互传感器设备和软件系统来实现。当前一个反复被提及和研讨的词汇是"AI"即人工智能(Artificial Intelligence),人工智能所强调的语言识别、图像识别、自然语言处理等技术形式,旨在帮助机器能够像人类一样思考,最终更好地与人类交流并服务于人类。智慧服务中的深度交互数据获取可以借鉴"AI"式的人机互动体验,使用户贡献数据得到深度挖掘的同时也能获得更优良的服务回报。为此,基于现有人机交互设备,创建用于深度交互数据获取功能。

在应用非常普及的移动终端,如手机、iPad、平板电脑、Kindle 上,设置文本对象涂抹功能(highlight),可让用户设置关注、批注和转发等功能互动操控,见图2。当用户阅读某一文献时,对文献中的表述某一观点的词或句子感兴趣,便可通过涂抹、批注和转发留下标记。对设备系统而言,这些文字内容的位置(篇章、页码、段落、行号)是完全掌控的,数据采集软件可以把用户的这些标记行为数据记录下来,作为深度交互数据而保存,这些数据再与文献资源和用户的基本数据关联,实现个性化的推送等智慧服务。

此外,一种能有效进行交互数据深度获取的方法是利用眼动仪。这种起初在心理学研究中被广泛运用的仪器,现在有望有效地解决深度数据交互问题,随着

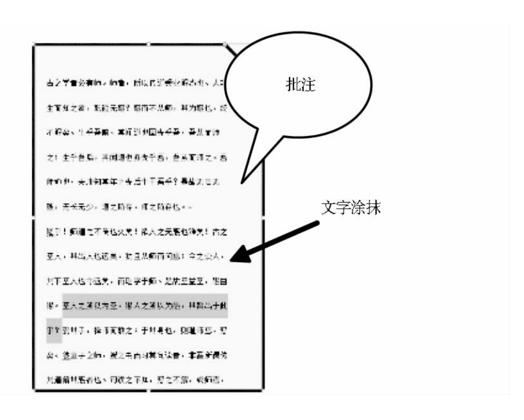


图 2 移动终端阅读的深度交互数据获取

眼动捕获传感器精度的提高,物理尺寸的微型化,未来有可能安装在移动设备上。

眼动仪通过对人眼球的运动轨迹的记录,观测用户的阅读顺序和视觉重点,深入地挖掘用户的阅读过程中所产生的数据[8],如图 3 所示。或者采用文字识别系统记录用户数据,目前文字识别技术趋于成熟,一大批文字识别处理的 APP 进入市场,使得即时、线上的文字处理和分享成为可能。以及其他类似于语音识别问答系统、活动图像交互判定等,都能为实现搭建图书馆与用户之间的深度的数据交互"桥梁"。

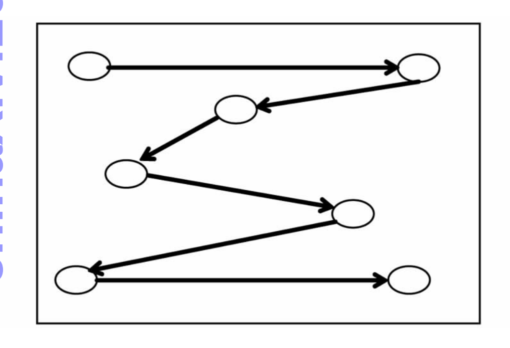


图 3 放置式眼动仪与阅读中的眼动轨迹

在眼动仪精度一定的情况下,为了提高深度交互数据获取的准确性,比如精确到某些词汇,这时可以通过调整屏幕阅读对象的几何大小来实现,如图 4 所示:

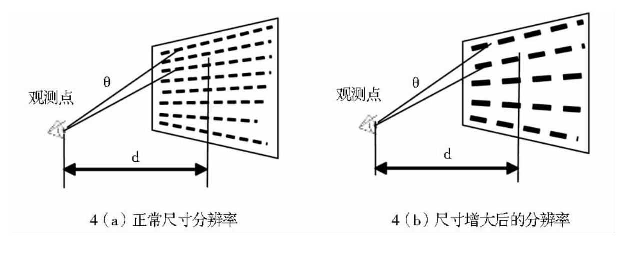


图 4 调整阅览对象尺寸提高分辨率

图 4(a)中,在观测者的分辨极限夹角 θ 情况下,设备只能捕获到三行文字的聚焦区,这样很难获得眼睛看到的是哪行文字,如果将文字字号调大到图 4(b)情况,在同样 θ 情况下,则可以分辨出观测者看到的是哪一行文字。

通过新设备、新技术的引入,图书馆的交互数据获取深度(D_d)可以大大提升。可以通过改良设备,调整阅读区域分辨率,提高数据获取准确度和深度。在对读者的阅读习惯、兴趣、方式等数据的细致鉴别和分析后的服务深度将比一般的交互性服务更加贴心和个性化,图书馆的智慧服务或许可以走向一个更智能、更具科技感的时代。

6 结语

智慧服务是现有的数字图书馆发展到智慧图书馆的过渡阶段而需要研究的问题。随着信息技术在图书馆的不断应用,图书馆服务模式将会发生重大变革。数据作为智慧时代的基础资源和驱动源动力,将被广泛重视,教育部近期批准一些高校设立数据科学专业就是最好的证明。本文揭示的现阶段高校图书馆的智慧服务内涵需要通过不断的理论和实践研究加以丰富和深入,智慧服务赖以实现的数据获取方式方法也有待于进一步实验量化。相信在不远的未来,数据获取技术的应用会让读者感受到图书馆智慧服务的魅力,也会推动智慧图书馆的实现。

参考文献:

[1] 周承聪,刘越强. 基于文献计量的国内智慧图书馆研究现状评析[J]. 新世纪图书馆,2017(10):89-96.
[2] 严栋. 基于物联网的智慧图书馆[J]. 图书馆学刊,2010,32(7):8-10.
[3] BUSCEMA M, TERZI S, MAURELLI G, et al. The smart library architecture of an orientation portal [J]. Quality & quantity, 2006, 40(6):911-933.
[4] 陈臣. 基于大数据的图书馆个性化智慧服务体系构建[J]. 情报资料工作,2013(6):75-79.
[5] 曾子明,陈贝贝. 融合情境的智慧图书馆个性化服务研究[J]. 图书馆论坛,2016,36(2):57-63.
[6] 陈远,许亮. 面向用户泛在智慧服务的智慧图书馆构建[J]. 图书馆杂志,2015,34(8):4-9.
[7] 刘宝瑞,沈苏阳. 用户体验视阈下的智慧图书馆研究[J]. 图书馆学研究,2017(6):43-47.
[8] 许鸿飞,高文君. 交互设计中眼动仪的可用性研究[J]. 科技传播,2017,9(22):145-146.
[9] 黄亚强. 以机器学习为基础的人工智能[J]. 电子技术与软件工程,2018(8):257.

李玉海，朱泽，夏红玉. 图书馆智慧服务中的数据获取[J]. 图书情报工作,2019,63(1):30－35.

作者贡献说明：

李玉海：研究思路设计、研究资料收集及论文审核定稿；

朱泽：论文初稿撰写及研究资料补充收集；

夏红玉：论文初稿修改及资料补充收集。

Data Acquisition in Library Intelligent Service

Li Yuhai　Zhu Ze　Xia Hongyu

School of Information Management, Central China Normal University, Wuhan 430079

Abstract: [**Purpose/significance**] This paper defines the connotation of library wisdom service, constructs the data model that supporting intelligent service, and discusses its methods and processes to promote the realization of intelligent services. [**Method/process**] This paper explains the meaning of the intelligent service in library, puts forward the relationship model between intelligence degree and interactivity, frequency, response time, depth of data perception of library services, discusses the influence of different data on intelligent service, and gives different ways of data acquisition. [**Result/conclusion**] Data is a necessary support for modern library to realize intelligent service. Libraries must rely on sufficient skills and patience to mine potential data of users, and realize rapid, in-depth and flexible intelligent services based on corresponding data interaction.

Keywords: smart library　intelligent service　data acquisition　data interaction